Problem 1

October 4, 2023 1:14 AM

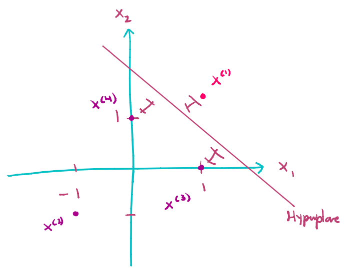
Dataset:

$$x^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad x^{(2)} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad x^{(3)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad x^{(4)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Positive, Negative, Negative, Negative

Or: $y^{(1)} = 1, \; y^{(2)} = y^{(3)} = y^{(4)} = -1$

1.



2. Support Vectors:

$x^{(1)}, x^{(3)}$ & $x^{(4)}$

3. Opt. Problem: $\min_{\vec{w}, b} \frac{1}{2} \|\vec{w}\|^2$ s.t. $\forall j \quad (\vec{w} \cdot \vec{x}^{(j)} + \vec{b}) y^{(j)} \geq 1$

Write down the Lagrangian:

$$\mathcal{L}(\vec{w}, \lambda) = f_0(\bar{x}) + \sum_{i=1}^{m} \lambda_i f_i(\bar{x})$$

$$f_0(\bar{x}) = \frac{1}{2} \|\vec{w}\|^2$$

$$f_j(\bar{x}) = j(\vec{w} \cdot \vec{x}^{(j)} + b) y^{(j)} - 1 \geq 0$$

$$\therefore \quad \mathcal{L}(\bar{w}, b, \bar{\alpha}) = \frac{1}{2} \|\vec{w}\|^2 - \sum_{j=1}^{m} a_j \left((\vec{w} \cdot \vec{x}^{(j)} + b) y^{(j)} - 1 \right)$$

4. $$\min_{\vec{w}, b} \max_{\bar{\alpha}} \mathcal{L}(\vec{w}, b, \bar{\alpha})$$

with Slater's Condition

$$\max_{\bar{\alpha}} \min_{\vec{w}, b} \mathcal{L}(\vec{w}, b, \vec{\alpha})$$

Solve min problem:

$$\max_{\vec{\alpha}} \min_{\vec{w},b} \frac{1}{2}\|\vec{w}\|^2 - \sum_{j=1}^{m} \alpha_j\left((\vec{w}\cdot\vec{x}^{(j)} + b)\, y^{(j)} - 1\right)$$

$$= \max_{\vec{\alpha}} \min_{\vec{w},b} \frac{1}{2}\|\vec{w}\|^2 - \sum_{j=1}^{m} \alpha_j \;\vec{w}\cdot\vec{x}^{(j)} y^{(j)} + \alpha_j \cdot b \cdot y^{(j)} - \alpha_j$$

$\hookrightarrow \frac{\partial}{\partial \vec{w}} \frac{1}{2}\|\vec{w}\|^2 = \vec{w}$

$$\frac{\partial \mathcal{L}}{\partial \vec{w}} = \vec{w} - \sum_{j=1}^{m} \alpha_j \vec{x}^{(j)} y^{j}$$

$$0 = \vec{w} - \sum_{j=1}^{m} \alpha_j x^{(j)} y^{(j)}$$

$$\boxed{\vec{w} = \sum_{j=1}^{m} \alpha_j \vec{x}^{(j)} y^{(j)}}$$

$$\frac{\partial \mathcal{L}}{\partial b} = \sum_{j=1}^{m} \alpha_j \cdot y^{(j)}$$

$$\boxed{0 = \sum_{j=1}^{m} \alpha_j y^{(j)}}$$

$\therefore \vec{A} = \vec{w}$, $B_j = a_j$, $C_j = y^j$, $\vec{D}_j = \vec{x}^{(j)}$

$E_j = \alpha_j$ $F_j = y^{(j)}$

5. Simplify:

$$\mathcal{L}(\vec{w},b,\alpha) = \frac{1}{2}\|\vec{w}\|^2 - \sum_{j=1}^{m} \alpha_j\left((\vec{w}\cdot\vec{x}^{(j)} + b)\, y^{(j)} - 1\right.$$

$$= \frac{1}{2}w^T w + \sum_{j=1}^{m} \alpha_j - \sum_{j=1}^{m} \alpha_j \vec{w}^T x^{(j)} y^{(j)} - \underbrace{\sum_{j=1}^{m} \alpha_j y^{(j)} b}$$

From above,
$\sum_{j=1}^{m} \alpha_j y^{(j)} = 0$
$\therefore$ whole thing $= 0$

$$= \frac{1}{2}w^T w + \sum_{j=1}^{m} \alpha_j - \sum_{j=1}^{m} \alpha_j \vec{w}^T \vec{x}^{(j)} y^{(j)}$$

$$= \sum_{j=1}^{m} \alpha_j + w^T\left(\frac{1}{2} w - \sum_{j=1}^{m} \alpha_j \vec{x}^{(j)} y^{(j)}\right.$$

From above

$$= \sum_{j=1}^{m} \alpha_j + w^T\left(\frac{1}{2}\sum_{j=1}^{m} \alpha_j \vec{x}^{(j)} y^{(j)} - \sum_{j=1}^{m} \alpha_j \vec{x}^{(j)} y^{(j)}\right)$$

From above

$$= \sum_{j=1}^{m} \alpha_j + w^T \left(\frac{1}{2} \sum_{j=1}^{m} \alpha_j \vec{x}^{(j)} y^{(j)} - \sum_{j=1}^{m} \alpha_j \vec{x}^{(j)} y^{(j)} \right)$$

$$= \sum_{j=1}^{m} \alpha_j + w^T \left(-\frac{1}{2} \sum_{j=1}^{m} \alpha_j \vec{x}^{(j)} y^{(j)} \right)$$

$$W = \sum_{j=1}^{m} \alpha_j y^{(j)} \vec{x}^{(j)} \rightarrow \text{Vector}$$

$$w^T = \sum_{h=1}^{n} a_h y^{(h)} \vec{x}^{T(h)}$$

$$= \sum_{j=1}^{m} \alpha_j - \frac{1}{2} \sum_{j=1}^{m} \sum_{h=1}^{n} \alpha_j \alpha_h y^{(j)} y^{(h)} \vec{x}^{T(h)} \vec{x}^{(j)}$$

$$\therefore \max_{\vec{a}} \min_{\vec{w}, b} \mathcal{L}(\vec{w}, b, \vec{a}) = \max_{\vec{a}} \sum_{j=1}^{m} \alpha_j - \frac{1}{2} \sum_{j=1}^{m} \sum_{h=1}^{n} \alpha_j \alpha_h y^{(j)} y^{(h)} (\vec{x}^{(h)})^T \vec{x}^{(j)}$$

$A_j = \alpha_j$, $B_j = \alpha_j$, $C_h = \alpha_{k}$, $D_j = y^{(j)}$, $E_h = y^{(h)}$

$\vec{F} = (x^{(h)})^T \qquad \vec{G} = x^{(j)}$

6: From observing the dataset that $x^{(2)} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ is not a support vector since its not close to the decision boundary. Therefore $a^{(2)} = 0$

7.

From Question 5.

$$0 = \sum_{j=1}^{m} a_j y^{(j)}$$

$$0 = a_1 y^1 + a_2 y^2 + a_3 y^3 + a_4 y^4$$

$$0 = a_1 (1) + 0(-1) + a_3(-1) + a_4(-1)$$

$$0 = a_1 - a_3 - a_4$$

$$a_1 = a_3 + a_4$$

8. Solve Dual Problem

$$\max_{\vec{a}} \min_{\vec{w},b} \mathcal{L}(\vec{w}, b, \vec{a}) = \max_{\vec{a}} \sum_{j=1}^{m} \alpha_j - \frac{1}{2} \sum_{j=1}^{m} \sum_{h=1}^{n} \alpha_j \alpha_h y^{(j)} y^{(h)} (\vec{x}^{(h)})^T \vec{x}^{(j)}$$

s.t. $\sum_{j=1}^{m} a_j y_j = 0$ & $a_j \geq 0$

$a_1 = a_3 + a_4$

$$= \max_{\vec{a}} \sum_{j=1}^{m} \alpha_j - \frac{1}{2} \sum_{j=1}^{m} \sum_{h=1}^{n} \alpha_j \alpha_h y^{(j)} y^{(h)} \langle x^{(h)} \cdot x^{(j)} \rangle$$

$$= \max_{\vec{a}} (a_1 + a_3 + a_4) - \frac{1}{2} \Big[\Big(a_1 a_1 (1)(1) [1\ 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix}\Big) + \Big(a_1 a_3 (1)(-1) [1\ 1] \begin{bmatrix} 1 \\ 0 \end{bmatrix}\Big) + \Big(a_1 a_4 (1)(-1) [1\ 1] \begin{bmatrix} 0 \\ 1 \end{bmatrix}\Big)$$

$$+ (a_3 a_1)(-1)(1)[1\ 0] \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \Big(a_3 a_3 (-1)(-1) [1\ 0] \begin{bmatrix} 1 \\ 0 \end{bmatrix}\Big) + \Big(a_3 a_4 \cancel{(-1)(-1)[1\ 0]} \begin{bmatrix} 0 \\ 1 \end{bmatrix}\Big) = 0$$

$$+ \Big(a_4 a_1 (-1)(1) [0\ 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix}\Big) + \Big(\cancel{a_4 a_3 (-1)(-1)} [0\ 1] \begin{bmatrix} 1 \\ 0 \end{bmatrix}\Big) + (a_4 a_4)(-1)(-1)[0\ 1] \begin{bmatrix} 0 \\ 1 \end{bmatrix}\Big)$$

$$= \max_{\vec{a}} (a_1 + a_3 + a_4) - \frac{1}{2} \left[2a_1^2 - a_1 a_3 - a_1 a_4 - a_3 a_1 + a_3^2 + \cancel{a_3 a_4} - a_4 a_1 + \cancel{a_4 a_3} + a_4^2 \right]$$

$$= \max_{\vec{a}} \left(a_1 + a_3 + a_4 - a_1^2 + a_1 a_3 + a_1 a_4 - \frac{a_3^2}{2} - \frac{a_4^2}{2} \right)$$

$$= \max_{\vec{a}} \left(-a_1^2 - \frac{a_3^2}{2} - \frac{a_4^2}{2} + a_1 a_3 + a_1 a_4 + a_1 + a_3 + a_4 \right)$$

Use $a_1 = a_3 + a_4$ to eliminate a variable

$$= \max_{\vec{a}} \left(-(a_3 + a_4)^2 - \frac{a_3^2}{2} - \frac{a_4^2}{2} + (a_3 + a_4) a_3 + (a_3 + a_4) a_4 + (a_3 + a_4) + a_3 + a_4 \right)$$

$$= \max_{\vec{a}} \left(-(a_3^2 + 2a_3 a_4 + a_4^2) - \frac{a_3^2}{2} - \frac{a_4^2}{2} + a_3^2 + a_4 a_3 + a_3 a_4 + a_4^2 + 2a_3 + 2a_4 \right)$$

$$= \max_{\vec{a}} \Big(-\cancel{a_3^2} - \cancel{2a_3 a_4} - \cancel{a_4^2} - \frac{a_3^2}{2} - \frac{a_4^2}{2} + \cancel{a_3^2} + \cancel{a_3 a_4} + \cancel{a_3 a_4} + \cancel{a_4^2} + 2a_3 + 2a_4$$

$$= \max_{\vec{a}} \left(-\cancel{a_3^2} - \cancel{2a_3a_4} - \cancel{a_4^2} - \frac{a_3^2}{2} - \frac{a_4^2}{2} + \cancel{a_3^2} + \cancel{a_3a_4} + \cancel{a_3a_4} + \cancel{a_4^2} + 2a_3 + 2a_4 \right.$$

$$= \max_{\vec{a}} \left(-\frac{a_3^2}{2} - \frac{a_4^2}{2} + 2a_3 + 2a_4 \right.$$

$$\frac{d\mathcal{L}}{d\vec{a}} = \begin{bmatrix} -a_3 + 2 & -a_4 + 2 \end{bmatrix}$$

$$0 = -a_3 + 2 \qquad 0 = -a_4 + 2$$

$$a_3 = 2 \qquad a_4 = 2$$

$$a_1 = a_3 + a_4 = 2 + 2$$

$$a_1 = 4$$

$$\vec{a} = \begin{bmatrix} 4 & 0 & 2 & 2 \end{bmatrix}$$

9. Solve for $\vec{w}$

$$\vec{w} = \sum_{j=1}^{m} a_j \vec{x}^{(j)} y^{(j)}$$

$$\vec{w} = 4\begin{bmatrix} 1 \\ 1 \end{bmatrix}(1) + 0\begin{bmatrix} -1 \\ -1 \end{bmatrix}(-1) + 2\begin{bmatrix} 1 \\ 0 \end{bmatrix}(-1) + 2\begin{bmatrix} 0 \\ 1 \end{bmatrix}(-1)$$

$$= \begin{bmatrix} 4 + 0 - 2 + 0 \\ 4 + 0 + 0 - 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

10. Find b using constraints in eqn 1:

$$(w^T \vec{x}^{(j)} + b) = y^{(j)}$$

$$b = y^{(j)} - w^T x^{(j)}$$

For $j=1$

$$b = 1 - [2\ 2]\begin{bmatrix}1\\1\end{bmatrix}$$
$$= 1 - 4$$
$$= -3$$

For $j=3$

$$b = -1 - [2\ 2]\begin{bmatrix}1\\0\end{bmatrix}$$
$$= -1 - 2$$
$$= -3$$

For $j=4$

$$b = -1 - [2\ 2]\begin{bmatrix}0\\1\end{bmatrix}$$
$$= -3$$

$\therefore\ b = -3$

II. Hyperplane: $y = \vec{w}\,\vec{x} + b$

$$y = [2\ 2]^T \vec{x} - 3$$

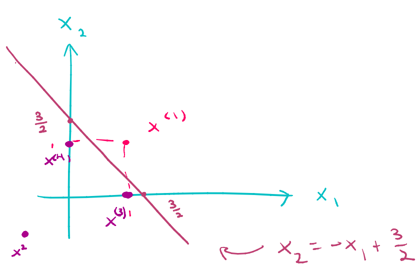
$$y = 2x_1 + 2x_2 - 3$$

Find relationship between x_2 & x_1

$$0 = 2x_1 + 2x_2 - 3$$

$$-2x_2 = 2x_1 - 3$$

$$x_2 = -x_1 + \frac{3}{2}$$



$\therefore$ This is the correct hyperplane based on the guess in part I